

**ДАНИЛОВ Г. В., КАЧАН О. В., БОРИСОВА Н. К.**  
**ОБ ОЦЕНКЕ ЦЕНТРОИДНОСТИ ПОИСКОВЫХ ДЕРЕВЬЕВ**  
*УДК 519.1, ВАК 05.13.18, ГРНТИ 28.29.51*

Об оценке центроидности поисковых  
деревьев

Evaluation of search trees centroidness

Г. В. Данилов<sup>1</sup>, О. В. Качан<sup>2</sup>,  
Н. К. Борисова<sup>1</sup>

G. V. Danilov<sup>1</sup>, O. V. Kachan<sup>2</sup>,  
N. K. Borisova<sup>1</sup>

<sup>1</sup>Ухтинский государственный  
технический университет, г. Ухта

<sup>1</sup>Ukhta state technical University,  
Ukhta

<sup>2</sup>Институт психологического  
консультирования «Новый век»,  
г. Санкт-Петербург

<sup>2</sup>Institute of psychological counseling  
"Novyy Vek", Saint Petersburg

*В статье описан подход к измерению степени центроидности поисковых и иных деревьев (связных графов без циклов). Доказано, что полученная мера принимает наименьшее возможное значение для дерева типа «ствол» и наибольшее – для дерева типа «куст».*

*The article describes the approach to measuring the degree of centroidness search and other trees (connected graphs without cycles). It is proved that the measure takes the smallest possible value for tree type "trunk" and the highest for wood type "Bush".*

**Ключевые слова:** связный граф, дерево, мера центроидности

**Keywords:** connected graph, tree, as centroidness

## **Введение**

Любая система, будь то физическая, информационная, экономическая, управленческая и т. д., обладает определённой структурой. Для математического моделирования структур обычно применяют такие объекты, как графы [1], представляющие собой множество  $V$  вершин и множество неупорядоченных пар элементов из  $V$ , называемых ребрами. Граф называется связным, если для любых его двух вершин существует маршрут, соединяющий эти вершины.

Наиболее простыми и часто встречающимися в приложениях графами являются деревья – связные графы, не содержащие циклов (рис. 1).

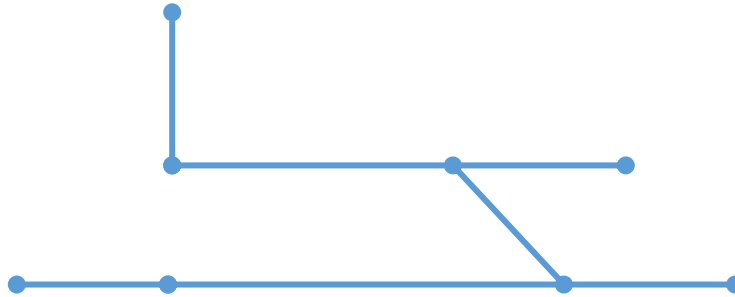


Рисунок 1. Пример дерева с 8 вершинами

Степенью вершины графа называется число инцидентных ей ребер. Для любого дерева справедливы следующие соотношения:

$$m = n - 1$$

$$\sum_{i=1}^n X_i = 2(n - 1) \quad (1)$$

количество висячих вершин не менее двух,

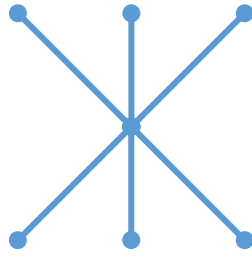
где  $n$  – число вершин,  $m$  – число ребер,  $X_i$  – степень  $i$ -й вершины ( $i = 1, 2, \dots, n$ ), «висячая» вершина – это вершина степени 1.

Во многих приложениях вершины, имеющие достаточно высокую степень, представляют значительный практический интерес как возможные центры кристаллизации, центры притяжения, источники или стоки, транспортные узлы, центроиды в поисковых системах и т. п. Подмножества таких вершин дерева в совокупности определяют его «ветвистость», или «центроидность».

### Количественная мера центроидности дерева

В связи с этим возникает необходимость ввести какую-то количественную меру центроидности дерева, которую обсудим пока на эвристическом уровне. Будем считать, что дерево типа «ствол» (рис. 2) является наименее ветвистым, а дерево типа «куст» (рис. 3) – наиболее ветвистым, и, следовательно, любая мера центроидности должна принимать наименьшее значение для ствола и наибольшее значение – для куста.

Рисунок 2. Дерево типа «ствол»  $n = 7$

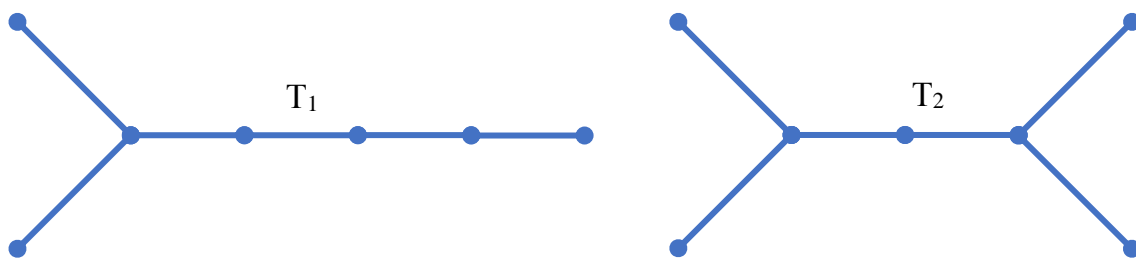
Рисунок 3. Дерево типа «куст»  $n = 7$ 

Первое, что приходит в голову, это ввести в качестве меры центроидности наибольшую степень вершины:

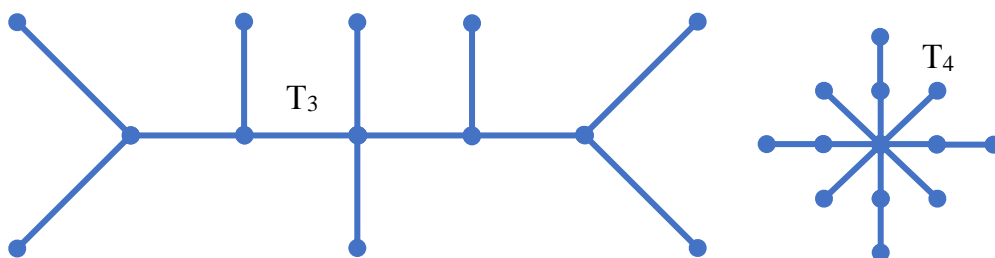
$$X_{\max} = \max_{v \in V} \{X(v)\} \quad (2)$$

Для ствола она принимает наименьшее значение, равное 2, а для куста наибольшее, равное  $n - 1$ .

Однако следующий пример (рис. 4) показывает, что такая мера не годится, поскольку оба дерева имеют одно и то же  $X_{\max} = 3$ , хотя дерево  $T_2$  интуитивно более ветвистое, чем  $T_1$ .

Рисунок 4. Пример двух деревьев ( $n = 7$ ) разной ветвистости, но имеющих одинаковые значения  $X_{\max} = 3$ 

Ещё одной мерой центроидности дерева могло бы быть число висячих вершин, которое также принимает наименьшее значение для ствола и наибольшее для куста. Но и эта мера не всегда соответствует нашим представлениям о ветвистости дерева. Например, на рисунке 5 при одинаковом числе висячих вершин дерево  $T_4$  представляется более ветвистым, чем  $T_3$ .

Рисунок 5. Пример двух деревьев ( $n = 13$ ) разной ветвистости, но имеющих одинаковое число висячих вершин (8)

Все эти примеры и контрпримеры наводят на мысль, что надо ввести такую меру центроидности, которая бы учитывала степени всех вершин дерева, а не отдельных его классов.

Рассмотрим упорядоченный по неубыванию перечень степеней всех  $n$  вершин дерева  $a = (X_1, X_2, \dots, X_n)$ ,  $X_1 \leq X_2 \leq \dots \leq X_n$ . Его можно рассматривать как точку в  $n$ -мерном пространстве степеней. Тогда множество всех деревьев с  $n$  вершинами разобьётся на классы, каждый из которых будет представлен определённой точкой, лежащей в гиперплоскости (1) в пространстве степеней, так что деревья, принадлежащие одному и тому же классу, будут неразличимы, т. е. принадлежность к данному классу будет для них топологическим инвариантом. Так, деревья  $T_2$  и  $T_5$  (рис. 6) будут представлены одной и той же точкой  $(1, 1, 1, 1, 2, 3, 3)$  пространства и в этом смысле неразличимы, хотя они и не являются изоморфными.

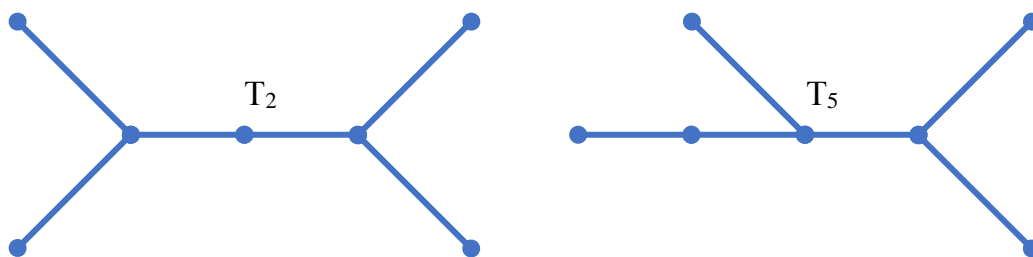


Рисунок 6. Пример двух различных деревьев ( $n = 7$ ), неразличимых в пространстве степеней

Теперь, считая, что пространство степеней наделено евклидовой метрикой, введем в качестве меры центроидности данного класса деревьев квадрат расстояния этого класса от начала координат:

$$C = \sum_{i=1}^n X_i^2 \quad (3)$$

Докажем, что эта мера (назовём ее «кроной») принимает наименьшее возможное значение для дерева типа «ствол» (Предложение 1) и наибольшее – для дерева типа «куст» (Предложение 2).

### Доказательство Предложения 1

В произвольном, отличном от ствола дереве, имеющем  $n$  вершин ( $n > 3$ ), Выделим две какие-нибудь висячие вершины. Не умаляя общности, можно считать, что они имеют номера  $n-1$  и  $n$  соответственно. Теперь осуществим следующую пошаговую процедуру.

Вначале удалим какую-нибудь висячую вершину (отличную от  $(n-1)^{\text{й}}$  и  $n^{\text{й}}$ ) вместе с инцидентным ей ребром и этот «тандем» присоединим к  $n^{\text{й}}$  висячей, так что степень последней станет равной двум. Затем удалим следующую висячую (отличную от  $(n-1)^{\text{й}}$ ) вместе со своим ребром, которое присоединим к только что «переселённой» висячей вершине. Вообще, на  $j^{\text{м}}$  шаге в дереве, образо-

ванном после  $(j-1)^{го}$  шага, удаляем произвольную, отличную от  $(n-1)^{й}$  вершину вместе с её ребром, которое присоединяем к висячей вершине, «переселённой» на  $(j-1)^{м}$  шаге.

Очевидно, за конечное число шагов все вершины будут иметь степень 2, за исключением  $(n-1)^{й}$  и последней переселившейся, которые будут висячими. Таким образом, описанная процедура всегда приводит к дереву типа ствол.

С другой стороны, на произвольном  $i^м$  шаге крона дерева  $C$  получает приращение  $C_i$ , равное:

$$\Delta C_i = (k-1)^2 - k^2 + 2^2 - 1^2 = 2(2-k),$$

где  $k$  – степень вершины, из которой на указанное шаге выдёргивается для переселения ребро; поскольку  $k = 2$ , то  $C_i = 0$ . Но в связи с тем, что всякое дерево, отличное от ствола, имеет хотя бы одну вершину степени больше 2 и из этой вершины на каком-то шаге (допустим  $1^м$ ) обязательно будет удалено ребро, то но крайней мере на этом шаге  $C_i < 0$ , и следовательно, суммарное приращение кроны будет отрицательным.

Поэтому ствол имеет наименьшую по сравнению с другими деревьями крону, и Предложение 1 доказано.

### Доказательство Предложения 2

В произвольном, отличном от куста дереве, имеющем  $n$  вершин ( $n > 3$ ), выделим вершину, обладающую наивысшей степенью. (Если таких вершин несколько, то выбираем любую из них).

Теперь осуществим следующую пошаговую процедуру.

На первом шаге удалим любую висячую вершину вместе с инцидентным ей ребром не соединённым с выделенной вершиной, и присоединим к последней этот «тандем». В результате выделенная вершина будет иметь наивысшую степень по сравнению со всеми другими вершинами. Назовём ее лидером.

Вообще, на  $j^м$  шаге очередную висячую вершину вместе с инцидентным ей ребром, не соединённым с лидером, отсоединяем и этот «тандем» присоединяем к лидеру.

Очевидно, за конечное число шагов все вершины (за исключением лидера) окажутся висячими и притом соединёнными с единственной вершиной - лидером.

Следовательно, описанная процедура всегда приводит к дереву типа куст.

С другой стороны, на каждом  $i^м$  шаге этой процедуры крона дерева  $C$  получает приращение  $C_i$ , равное:

$$\Delta C_i = (k_i - 1)^2 - k_i^2 + (l - 1)^2 - l^2 = 2(k_i - l + 1),$$

где  $k_i$  – степень лидера перед  $i^м$  шагом;  $l$  – степень вершины, из которой на данном шаге выдёргивается ребро.

Поскольку  $K_i = 1$ , то  $C_i > 0$ . Поэтому куст обладает наибольшей кроной по сравнению с другими деревьями, и Предложение 2 доказано.

Отметим, что в приведённых на рисунках 4 и 5 примерах крона принимает большее значение для тех деревьев, которые как раз и представлялись нам более ветвистыми:  $C(T_2) = 26 > 24 = C(T_3)$  и  $C(T_4) = 88 > 60 = C(T_3)$ .

После соответствующей нормировки получаем выражение для количественной меры центроидности дерева  $C^*$ :

$$C^* = \frac{4}{n-3} \cdot (C' - 1) \quad (0 \leq C^* \leq 1)$$

где  $C' = \frac{C-2}{4(n-2)}$ , которая принимает значение 0 для ствола и значение 1 для куста.

### **Заключение**

Следующим возможным (но более трудным) шагом было бы отыскание некоторой статистики как функции кроны, которая на данном уровне значимости позволяла бы судить о значимо «богатой» или значимо «бедной» кроне случайно взятого класса деревьев и на основе этого делать содержательные выводы относительно механизма, порождающего этот класс.

### **Список литературы**

1. Харари Ф. Теория графов. М. : Мир, 1973. 283 с.

### **List of references**

1. F. Harari. *Graph Theory*. Moscow : Peace, 1973. 283 p.